N° 82-003-X au catalogue
ISSN 1209-1375

Rapports sur la santé

Couplage de l'Enquête sur la santé dans les collectivités canadiennes et de la Base canadienne de données sur la mortalité : une source de données améliorée pour l'étude de la mortalité

par Claudia Sanmartin, Yves Decady, Richard Trudeau, Abel Dasylva, Michael Tjepkema, Philippe Finès, Rick Burnett, Nancy Ross et Douglas G. Manuel

Date de diffusion : le 21 décembre 2016

Statistique Canada Statistics Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à STATCAN.infostats-infostats.STATCAN@canada.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- Service de renseignements statistiques 1-800-263-1136
- Service national d'appareils de télécommunications pour les malentendants 1-800-363-7629
- Télécopieur 1-514-283-9350

Programme des services de dépôt

- Service de renseignements 1-800-635-7943
- Télécopieur 1-800-565-7757

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « Normes de service à la clientèle ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Signes conventionnels dans les tableaux

Les signes conventionnels suivants sont employés dans les publications de Statistique Canada :

.	indisponible pour toute période de référence
..	indisponible pour une période de référence précise
...	n'ayant pas lieu de figurer
0	zéro absolu ou valeur arrondie à zéro
0[s]	valeur arrondie à 0 (zéro) là où il y a une distinction importante entre le zéro absolu et la valeur arrondie
[p]	provisoire
[r]	révisé
x	confidentiel en vertu des dispositions de la *Loi sur la statistique*
[E]	à utiliser avec prudence
F	trop peu fiable pour être publié
*	valeur significativement différente de l'estimation pour la catégorie de référence (p<0,05)

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2016

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'entente de licence ouverte de Statistique Canada.

Une version HTML **est aussi disponible.**

This publication is also available in English.

Couplage de l'Enquête sur la santé dans les collectivités canadiennes et de la Base canadienne de données sur la mortalité : une source de données améliorée pour l'étude de la mortalité

par Claudia Sanmartin, Yves Decady, Richard Trudeau, Abel Dasylva, Michael Tjepkema, Philippe Finès, Rick Burnett, Nancy Ross et Douglas G. Manuel

Résumé

Contexte : La présente étude donne un aperçu du couplage qui a été effectué entre l'Enquête sur la santé dans les collectivités canadiennes (ESCC) et la Base canadienne de données sur la mortalité (BCDM) pour examiner les rapports entre les déterminants sociaux, les comportements en matière de santé et la mortalité dans la population à domicile.

Données et méthodes : Les données de l'Enquête sur la santé dans les collectivités canadiennes pour 2000-2001 à 2011 ont été couplées à celles de la BCDM pour les années 2000 à 2011, au moyen de méthodes probabilistes, sur la base d'identificateurs communs (noms, date de naissance, code postal et sexe) pour les participants admissibles (85 %; n = 614 774). Les enregistrements de la mortalité survenue du 1er janvier 2000 au 31 décembre 2011, inclusivement, pour les personnes âgées de 12 ans et plus étaient admissibles au couplage (n = 2,774 millions). Le couplage a été amélioré grâce aux données du Fichier de données fiscales sommaires historiques. L'évaluation de la qualité a pris la forme d'une validation interne et externe. Une analyse de survie de Cox (corrigée selon l'âge) a été menée pour estimer les rapports de risques (RR) associés à certains comportements en matière de santé.

Résultats : Dans l'ensemble, 5,3 % des participants à l'ESCC admissibles ont fait l'objet d'un couplage à un enregistrement de la mortalité; les taux de faux positifs et de faux négatifs ont été de 0,04 % et de 2,43 %, respectivement. Les taux de couplage étaient plus élevés chez les hommes (5,8 %) et les personnes de 75 ans et plus (20,2 %), ce qui rend compte des risques de mortalité connus. Des analyses de survie ont confirmé un risque de mortalité plus grand lié à l'usage élevé (RR 2,36; IC 1,84 - 2,89) et faible du tabac (RR 1,91; IC 1,52 - 2,33), comparativement à l'absence d'usage du tabac; à l'insuffisance pondérale et à l'obésité, comparativement au poids normal et à l'embonpoint; à la faible consommation de fruits et de légumes; et à l'absence d'activité physique.

Interprétation : Le couplage des données sur les comportements en matière de santé de l'ESCC aux données sur la mortalité de la BCDM permet de mieux comprendre les déterminants modifiables de la mortalité.

Mots-clés : Couplage de données, enquête sur la santé, mortalité, facteurs de risque.

Dans la plupart des pays industrialisés, les registres de la statistique de l'état civil et les enquêtes nationales sur la santé sont les pierres angulaires de la surveillance de la santé.

Les données sur la mortalité compilées par les registres de la statistique de l'état civil à des fins administratives peuvent être totalisées selon des caractéristiques démographiques de base (âge et sexe), la province et la cause du décès. Toutefois, on sait peu de choses concernant les caractéristiques socioéconomiques, culturelles ou linguistiques des personnes qui décèdent, ou encore au sujet de l'incidence du mode de vie et des facteurs sociaux sur le risque de mortalité.

Les enquêtes sur la santé, par contre, servent à recueillir des données concernant l'état de santé et les comportements en matière de santé des participants, ainsi que sur leurs caractéristiques socioéconomiques et culturelles. Toutefois, ces enquêtes sont habituellement transversales et aucune donnée n'est fournie concernant la santé des participants, outre la date de l'enquête.

Lorsque des enquêtes sur la santé et des données administratives complémentaires sont combinées grâce au couplage d'enregistrements[1-3], les rapports entre les déterminants sociaux et les résultats pour la santé, comme le décès, peuvent être analysés de façon approfondie[4-8]. Par exemple, au Canada, les données sur la mortalité ont été couplées aux résultats du recensement pour examiner les différences dans le risque de décès entre les groupes socioéconomiques[9,10]. Ces données couplées ont aussi été utilisées pour calculer les taux de mortalité de sous-populations, comme les immigrants et les groupes autochtones[11-13]. De même, les chercheurs d'autres pays ont utilisé les données couplées pour examiner la mortalité chez les immigrants, les Autochtones et les prisonniers[14-17]. Une étude repère en Ontario a utilisé les données couplées des enquêtes sur la santé et de la mortalité pour explorer les associations avec l'usage du tabac, un mauvais régime alimentaire, l'inactivité physique et un niveau élevé de stress. Les résultats ont démontré que 60 % des décès en Ontario étaient attribuables à ces facteurs de risque[18].

Le couplage décrit dans la présente étude regroupe des données de l'Enquête sur la santé dans les collectivités canadiennes et des données sur la mortalité de la Base canadienne de données sur la mortalité. Le présent article explique le processus de couplage des enregistrements et décrit les résultats concernant les associations entre les comportements en matière de santé et la mortalité dans un échantillon représentatif de Canadiens.

Auteurs : Claudia Sanmartin (Claudia.Sanmartin@canada.ca), Michael Tjepkema, Philippe Finès et Douglas G. Manuel travaillent à la Division de l'analyse de la santé, Yves Decady, à la Division de la statistique de la santé, Richard Trudeau, à la Division des enquêtes spéciales, et Abel Dasylva, à la Division des méthodes d'enquêtes auprès des ménages, de Statistique Canada. Rick Burnett est de Santé Canada et Nancy Ross est de l'Université McGill. Douglas G. Manuel est aussi chercheur à l'Institut de recherche de l'Hôpital d'Ottawa.

Données et méthodes

Enquête sur la santé dans les collectivités canadiennes

L'Enquête sur la santé dans les collectivités canadiennes (ESCC) est une enquête transversale qui a pour but de recueillir des données sur la santé, les comportements en matière de santé et l'utilisation des soins de santé de la population à domicile de 12 ans et plus ne vivant pas en établissement. L'enquête exclut les membres à temps plein des Forces canadiennes et les habitants des réserves et de certaines régions éloignées, qui représentent ensemble environ 4 % de la population cible. L'ESCC a été menée pour la première fois en 2000-2001 (cycle 1), et de nouveau en 2003 (cycle 2) et en 2005 (cycle 3), chaque fois à partir d'un échantillon d'environ 130 000 participants. À partir de 2007, l'enquête a été menée sur une base annuelle (taille d'échantillon de 65 000). Les taux de réponse ont varié de 69,8 % à 78,9 %. Des détails concernant la stratégie d'échantillonnage et le contenu sont accessibles dans le site Web de Statistique Canada (www.statcan.gc.ca)[19].

Les enregistrements de l'ESCC admissibles au couplage avec la Base canadienne de données sur la mortalité correspondaient aux participants qui avaient consenti à partager leurs données avec les ministères provinciaux et fédéral de la Santé, et à coupler leurs réponses avec des données administratives. Dans l'ensemble, 85,3 % (n = 614 774) des participants à l'ESCC étaient admissibles au couplage (tableau 1). Des poids d'échantillonnage spéciaux ont été créés pour les données couplées de l'ESCC pour tenir compte des personnes qui n'avaient pas consenti au partage et au couplage.

Tableau 1
Participants à l'Enquête sur la santé dans les collectivités canadiennes admissibles au couplage d'enregistrements avec la Base canadienne de données sur la mortalité, selon l'année d'enquête, 2000-2001 à 2011

Année de l'Enquête sur la santé dans les collectivités canadiennes (cycle)		Ont consenti au partage et au couplage	
	nombre	nombre	%
Total	**720 426**	**614 774**	**85,3**
2000-2001 (1,1)	131 535	117 837	89,6
2003-2004 (2,1)	135 573	112 851	83,2
2005-2006 (3,1)	132 947	113 880	85,7
2007	65 946	57 083	86,6
2008	66 013	55 592	84,2
2009	61 679	52 475	85,1
2010	63 191	52 198	82,6
2011	63 542	52 858	83,2

Source : Enquête sur la santé dans les collectivités canadiennes, 2000-2001 à 2011.

Base canadienne de données sur la mortalité

La Base canadienne de données sur la mortalité (BCDM) est un recensement de tous les décès enregistrés au Canada. Les décès sont déclarés à Statistique Canada par les registres provinciaux et territoriaux de la statistique de l'état civil. L'information fournie comprend la cause et la date du décès, les noms, la date de naissance et le code postal au moment du décès. Les données sur la cause de décès sont codées au moyen de la version de la *Classification internationale des maladies* (CIM) en vigueur au moment du décès. Les enregistrements de la BCDM admissibles au couplage correspondent aux décès de personnes de 12 ans et plus (n = 2,774 millions) survenus du 1er janvier 2000 au 31 décembre 2011 inclusivement.

Fichier de données fiscales sommaires historiques

Le Fichier de données fiscales sommaires historiques (FDFSH) est une compilation des fichiers annuels de l'impôt sur le revenu représentant des particuliers pour lesquels une déclaration de revenu a été produite au cours d'une année donnée. Le FDFSH renferme uniquement des données d'identification personnelles (noms, date de naissance, sexe et code postal) tirées du Fichier maître des particuliers de l'impôt sur le revenu T1, lequel ne comprend pas de données sur le revenu. Statistique Canada utilise le FDFSH pour faciliter le couplage des enregistrements, en particulier grâce aux autres données de couplage qui s'y trouvent (noms, code postal), ainsi que dans la résolution manuelle des couplages douteux.

Les données du FDFSH pour la période de 1996 à la fin de 2011, qui représentent environ 33,5 millions d'enregistrements, ont servi au couplage. Les données tirées du FDFSH et utilisées pendant le couplage ont été supprimées du fichier d'analyse final.

Couplage

Le couplage comprenait les trois étapes suivantes : 1) préparation des données; 2) couplage des enregistrements; 3) évaluation de la qualité. Le couplage a été approuvé par le Conseil exécutif de gestion de Statistique Canada[20]. L'utilisation des données couplées est régie par la Directive sur le couplage d'enregistrements[21].

Préparation des données

Les variables suivantes, qui ont été choisies sur la base des aspects communs aux ensembles de données, de la qualité des données et du pouvoir discriminant, ont servi au couplage : prénom, nom, date de naissance, sexe et code postal. Une réponse invalide fournie à l'une ou l'autre des variables était fixée à « données manquantes ».

Le taux de données manquantes pour les variables de couplage était habituellement inférieur à 3 %. Le nom saisi dans le cadre de l'ESCC faisait exception. En moyenne, lorsqu'on tenait compte de tous les cycles de l'enquête, 10 % des participants avaient des données manquantes pour le nom. Ce pourcentage allait de 4,1 % en 2000-2001 à 17,1 % en 2011. Par suite du couplage avec le FDFSH, le pourcentage de participants à l'ESCC dont le nom était manquant a diminué, pour passer à 2 %. Pour réduire le nombre de couplages manqués en raison d'erreurs d'orthographe, tous les noms (prénoms et noms) ont été convertis à leur forme phonétique au moyen du New York State Individual Intelligence System.

Le pouvoir discriminant des variables de couplage a été évalué au moyen des scores d'entropie de Shannon[22]. Des niveaux élevés d'entropie reflètent un grand pouvoir discriminant. Les vari-

ables comportant un plus grand nombre de valeurs distinctes (p. ex. le nom) ont un pouvoir discriminant et des niveaux d'entropie plus élevés comparativement aux variables comportant moins de valeurs distinctes (p. ex. le sexe)[23]. Les niveaux d'entropie parmi les variables de couplage allaient de 0,43 pour le prénom à 0,85 pour le code postal (données non présentées dans les tableaux). On a tenu compte de ces renseignements dans la stratégie de couplage d'enregistrements, aux fins d'attribution des poids de couplage initiaux.

Une seule cohorte de couplage a été créée en concaténant les variables de couplage de tous les cycles de l'ESCC. Cette approche est plus efficace et réduit la probabilité de faux couplages comparativement au couplage de chaque cycle séparément.

D'autres détails concernant le couplage d'enregistrements sont disponibles dans un document interne[24].

Couplage d'enregistrements

Le couplage d'enregistrements de l'ESCC et de la BCDM s'est fait en deux étapes. Tout d'abord, les enregistrements des participants à l'ESCC qui y étaient admissibles ont été couplés au FDFSH, en vue d'y joindre d'autres codes postaux et noms possibles. Ces renseignements ont servi à la deuxième étape, au moment du couplage de l'ESCC et de la BCDM. Les enregistrements de tous les participants à l'ESCC admissibles ont été couplés à la BCDM, peu importe si l'enregistrement était couplé au FDFSH.

Les couplages ont été effectués en G-Coup, selon la méthode de couplage probabiliste fondée sur la théorie du couplage d'enregistrements de Fellegi-Sunter[25]. G-Coup est un logiciel de couplage d'enregistrements généralisé tournant sous SAS, élaboré par Statistique Canada pour faciliter le couplage probabiliste à grande échelle[26].

L'ESCC a été couplée au FDFSH en fonction des variables suivantes : *prénom et nom, date de naissance, code postal* et *sexe*. L'ESCC a été couplée à la BCDM à partir des mêmes variables, présentes dans les deux ensembles de données. Dans le cas des enregistrements couplés au FDFSH, les variables suivantes ont également été utilisées : *autres codes postaux, autres noms (p. ex. nom de fille, nom du père), date du décès et date de l'entrevue*. On a déterminé que l'accès à d'autres codes postaux dans le FDFSH améliore les taux de couplage, particulièrement au fil du temps[27].

Les taux de couplage ont été calculés pour le couplage de l'ESCC et du FDFSH et pour celui de l'ESCC et de la BCDM. Les taux représentent le nombre de participants à l'ESCC dont les enregistrements ont fait l'objet d'un couplage divisé par le nombre de participants à l'ESCC qui étaient admissibles au couplage (non pondérés et pondérés).

Évaluation de la qualité

On a procédé à une validation interne et externe pour évaluer la précision du processus de couplage et l'adaptation des données à l'analyse.

Validation interne

On a calculé les taux de faux positifs et de faux négatifs. Un examen manuel (à l'aveugle) a été effectué relativement à un échantillon aléatoire de 1 000 paires sélectionnées uniformément dans huit strates de poids de couplage représentant la gamme complète de poids de couplage supérieurs au seuil (c.-à-d. les couplages), mais seulement une fourchette étroite de poids juste en deçà du seuil (les non-couplages). Cette approche est recommandée compte tenu du nombre important de non-couplages générés au moment de la création initiale des paires. Si ce nombre était pleinement représenté, cela entraînerait un taux de faux négatifs artificiellement faible[28].

Trois examinateurs ont procédé individuellement à un examen manuel des mêmes données que celles utilisées pour le couplage (noms, date de naissance, code postal et sexe). Parmi les autres renseignements dont on disposait aux fins de l'examen manuel figurait la date de décès, dans le FDFSH, des participants à l'ESCC ayant fait l'objet d'un couplage avec les données fiscales et qui étaient décédés. On a ensuite attribué le statut de couplage ou de non-couplage à chacune des paires.

Tableau 2
Nombre et pourcentage de participants à l'Enquête sur la santé dans les collectivités canadiennes (ESCC) qui ont fait l'objet d'un couplage au Fichier de données fiscales sommaires historiques (FDFSH) et à la Base canadienne de données sur la mortalité (BCDM), selon l'année d'enquête, 2000-2001 à 2011

Année de l'Enquête sur la santé dans les collectivités canadiennes	Couplage avec le FDFSH				Couplage avec la BCDM				Couplage avec la BCDM (participants à l'ESCC pour qui a eu lieu un couplage au FDFSH)			
	Non pondéré		Pondéré		Non pondéré		Pondéré		Non pondéré		Pondéré	
	nombre	%	nombre	%	nombre	%	nombre	%	nombre	%	nombre	%
Total	**517 725**	**84,2**	**23 015 391**	**82,8**	**32 889**	**5,3**	**.**	**.**	**31 274**	**6,0**	**.**	**.**
2000-2001 (1,1)	101 495	86,1	22 063 102	85,5	11 768	10,0	1 979 060	7,7	11 187	11,0	1 869 609	8,5
2003-2004 (2,1)	92 881	82,3	21 302 890	80,2	8 672	7,7	1 399 526	5,3	8 222	8,9	1 327 901	6,2
2005-2006 (3,1)	98 590	86,6	23 190 360	85,5	6 685	5,9	1 092 021	4,0	6 345	6,4	1 033 059	4,5
2007	49 096	86,0	23 725 215	85,2	2 302	4,0	763 537	2,7	2 211	4,5	723 362	3,0
2008	47 046	84,6	23 696 053	84,0	1 689	3,0	552 636	2,0	1 620	3,4	534 562	2,3
2009	43 723	83,3	23 483 307	82,2	992	1,9	316 691	1,1	951	2,2	302 384	1,3
2010	42 962	82,3	23 534 665	81,5	622	1,2	220 092	0,8	586	1,4	199 831	0,8
2011	41 932	79,3	23 127 535	79,2	159	0,3	63 076	0,2	152	0,4	57 034	0,2

. indisponible pour toute période de référence

Sources : Enquête sur la santé dans les collectivités canadiennes, 2000-2001 à 2011; Fichier de données fiscales sommaires historiques, 1996 à 2011; Base canadienne de données sur la mortalité, 2000 à 2011.

Les résultats de l'examen manuel ont été comparés à ceux de G-Coup, afin de calculer les estimations par strate des couplages faussement positifs et faussement négatifs[26,29]. Les taux globaux de faux positifs et de faux négatifs ont été estimés au moyen d'une fonction de pondération fondée sur la répartition de toutes les paires possibles. Même si on a eu recours à une répartition d'échantillon uniforme pour sélectionner les paires dans chaque strate de poids, les fractions d'échantillonnage différaient parce que certaines strates comprenaient davantage de paires que d'autres. Environ 75 % des paires appartenaient à la première strate, dans laquelle les poids de couplage étaient les plus faibles; moins de 1 % des paires appartenaient à la dernière strate (strate 8), dans laquelle les poids de couplage étaient les plus élevés.

Validation externe

Les taux de couplage avec la BCDM ont été calculés selon les caractéristiques de base des participants (p. ex. âge et sexe). On prévoyait que les tendances des taux de couplage refléteraient le risque différentiel de décès parmi les groupes les plus à risque, comme les hommes et les adultes les plus âgés[30,31].

Les données couplées ont également fait l'objet d'une validation externe, qui consistait à comparer la répartition des décès pour la cohorte issue du couplage avec celle calculée à partir de la BCDM uniquement, pour les personnes de 12 ans et plus. On a examiné les répartitions selon la région géographique, l'âge, le sexe et la principale cause de décès. On a calculé les taux de mortalité normalisés selon l'âge (TMNA) par année civile pour chaque cycle de l'ESCC, puis on a comparé ceux-ci aux statistiques officielles sur la mortalité[32].

Finalement, des modèles de survie à risques proportionnels de Cox (corrigés selon l'âge) ont servi à évaluer les répercussions de plusieurs facteurs de risque comportementaux sur la mortalité. L'analyse portait sur les participants à l'ESCC de 20 ans et plus ayant fait l'objet d'un couplage à partir des données de 2003 et 2005 de l'enquête. Le suivi de la mortalité s'est échelonné sur une période maximale de huit ans. À cause des résultats de la validation externe, la première année de suivi a été exclue de l'analyse.

Tableau 3
Nombre et pourcentage de participants à l'Enquête sur la santé dans les collectivités canadiennes (ESCC) de 12 ans et plus qui ont fait l'objet d'un couplage à la Base canadienne de données sur la mortalité (BCDM), selon la province ou le territoire, le sexe et le groupe d'âge, 2000-2001 à 2011

Province/territoire, sexe et groupe d'âge	Participants à l'ESCC pour qui a eu lieu un couplage à la BCDM			
	Non pondéré		Pondéré	
	nombre	%	nombre	%
Total	**32 889**	**5,3**	..	..
Province/territoire				
Terre-Neuve-et-Labrador	950	5,0	15 374	3,4
Île-du-Prince-Édouard	746	6,6	4 279	3,6
Nouvelle-Écosse	1 622	6,6	30 644	3,8
Nouveau-Brunswick	1 372	5,6	21 683	3,4
Québec	6 068	4,9	198 841	3,0
Ontario	10 382	5,3	291 177	2,7
Manitoba	2 228	6,2	32 065	3,4
Saskatchewan	2 413	6,6	28 913	3,6
Alberta	2 638	4,5	65 655	2,3
Colombie-Britannique	4 144	5,7	108 248	2,9
Yukon	122	3,0	532	1,9
Territoires du Nord-Ouest	118	2,4	587	1,7
Nunavut	86	2,5	327	2,0
Sexe				
Hommes	16 439	5,8	429 275	3,1
Femmes	16 450	5,0	369 054	2,6
Groupe d'âge				
12 à 24 ans	310	0,3	9 138	0,2
25 à 34 ans	356	0,4	10 804	0,2
35 à 44 ans	932	1,0	29 068	0,6
45 à 54 ans	2 242	2,4	73 536	1,5
55 à 64 ans	4 482	4,8	122 125	3,2
65 à 74 ans	8 233	11,4	191 224	8,0
75 ans et plus	16 334	27,0	362 432	20,5

.. indisponible pour une période de référence précise
Sources : Enquête sur la santé dans les collectivités canadiennes, 2000-2001 à 2011; Base canadienne de données sur la mortalité, 2000 à 2011.

Covariables

Les résultats du couplage sont présentés selon la province et le territoire, le sexe, et l'âge au moment de l'entrevue de l'ESCC. Les résultats de l'analyse de survie sont déclarés pour les cinq facteurs de risque suivants, au sens de l'ESCC : situation d'usage du tabac, indice de masse corporelle (IMC), consommation d'alcool, activité physique et consommation de fruits et de légumes. Les valeurs autodéclarées d'IMC ont été corrigées au moyen du facteur de correction recommandé[33]. On a utilisé les poids d'échantillonnage spéciaux dans les analyses pour tenir compte de l'exclusion des participants à l'ESCC qui n'avaient pas consenti au partage et au couplage de leurs données.

Respect de la vie privée des participants à l'enquête

Statistique Canada assure le respect de la vie privée des participants aux enquêtes au cours du processus de couplage et de l'utilisation subséquente des fichiers couplés. Seuls les employés participant directement au processus ont accès aux données d'identification uniques requises pour le couplage (comme les noms). Ils n'ont *pas* accès aux données liées à la santé. Une fois le couplage de données

terminé, on crée un fichier d'analyse, duquel les données d'identification sont supprimées. Les analystes ont par la suite accès au fichier anonymisé aux fins de validation et d'analyse.

Résultats

Couplage

Une grande majorité (84,2 %) des participants à l'ESCC qui étaient admissibles au couplage ont fait l'objet d'un appariement au FDFSH, les taux variant de 79,3 % à 86,6 % (tableau 2) sur la période s'étirant de la première à la dernière année d'enquête visée par l'étude. Les taux de couplage étaient faibles (78 %) pour la Colombie-Britannique et le Nunavut comparativement aux autres régions (données non présentées dans les tableaux). De même, les taux de couplage étaient faibles chez les participants de moins de 25 ans (74,3 %) ou de plus de 74 ans (79,4 %), et élevés chez les participants de 25 à 34 ans (89 %) (données non présentées dans les tableaux). Ces tendances selon la région et l'âge ont persisté tout au long de la période couvrant les années d'enquête (données non présentées).

Dans l'ensemble, 5,3 % des participants à l'ESCC qui étaient admissibles ont fait l'objet d'un appariement à la BCDM (tableau 2). Le taux de couplage à la BCDM était légèrement plus élevé parmi les participants à l'ESCC pour lesquels s'était produit un couplage au FDFSH (6,0 %). Comme on pouvait s'y attendre, compte tenu de la période de suivi plus courte, les taux de couplage étaient plus faibles dans le cas des années d'enquête récentes. Les participants des territoires étaient moins susceptibles de faire l'objet d'un couplage à la BCDM (2,4 % à 3,0 %) que ceux des autres régions.

Tableau 4
Répartition en pourcentage pondérée et non pondérée des décès parmi les personnes de 12 ans et plus relevés dans la cohorte formée par suite du couplage avec l'Enquête sur la santé dans les collectivités canadiennes (ESCC), comparativement à la Base canadienne de données sur la mortalité (BCDM), selon la province ou le territoire, le sexe, le groupe d'âge et la cause de décès, 2000-2001 à 2011

	Participants à l'ESCC pour qui a eu lieu un couplage à la BCDM		
	Non pondérée	Pondérée	BCDM
	%		
Province/territoire			
Terre-Neuve-et-Labrador	2,9	1,9	1,9
Île-du-Prince-Édouard	2,3	0,5	0,5
Nouvelle-Écosse	4,9	3,8	3,6
Nouveau-Brunswick	4,2	2,7	2,7
Québec	18,4	24,9	24,3
Ontario	31,6	36,5	37,0
Manitoba	6,8	4,0	4,3
Saskatchewan	7,3	3,6	3,9
Alberta	8,0	8,2	8,4
Colombie-Britannique	12,6	13,6	13,2
Yukon	0,4	0,1	0,1
Territoires du Nord-Ouest	0,4	0,1	0,1
Nunavut	0,3	0,0	0,0
Sexe			
Hommes	50,0	53,8	50,5
Femmes	50,0	46,2	49,5
Groupe d'âge			
12 à 24 ans	0,9	1,1	1,2
25 à 34 ans	1,1	1,4	1,2
35 à 44 ans	2,8	3,6	2,6
45 à 54 ans	6,8	9,2	6,0
55 à 64 ans	13,6	15,3	10,6
65 à 74 ans	25,0	24,0	17,7
75 ans et plus	49,0	45,4	60,7
Cause de décès			
Maladies du sang et des organes hématopoïétiques et certains troubles du système immunitaire	0,4	0,5	0,4
Cancers	33,9	36,1	30,2
Maladies de l'appareil circulatoire	31,0	28,9	31,2
Maladies de l'appareil digestif	4,0	3,8	3,9
Maladies endocriniennes, nutritionnelles et métaboliques	4,1	3,9	4,2
Causes externes de morbidité et de mortalité	5,5	5,7	6,4
Maladies de l'appareil génito-urinaire	2,3	2,4	2,2
Certaines maladies infectieuses et parasitaires	1,7	2,0	1,8
Malformations congénitales et anomalies chromosomiques	0,1	0,2	0,2
Troubles mentaux et du comportement	3,0	3,0	4,0
Maladies du système ostéo-articulaire, des muscles et du tissu conjonctif	0,7	0,7	0,7
Maladies du système nerveux	3,4	3,7	4,8
Maladies de l'appareil respiratoire	8,6	8,3	8,6
Maladies de la peau et du tissu cellulaire sous-cutané	0,2	0,1	0,2
Symptômes, signes et résultats anormaux d'examens cliniques et de laboratoire	1,0	0,9	1,1
Autres causes	0,0	0,0	0,0

Sources : Enquête sur la santé dans les collectivités canadiennes, 2000-2001; Base canadienne de données sur la mortalité.

Évaluation de la qualité

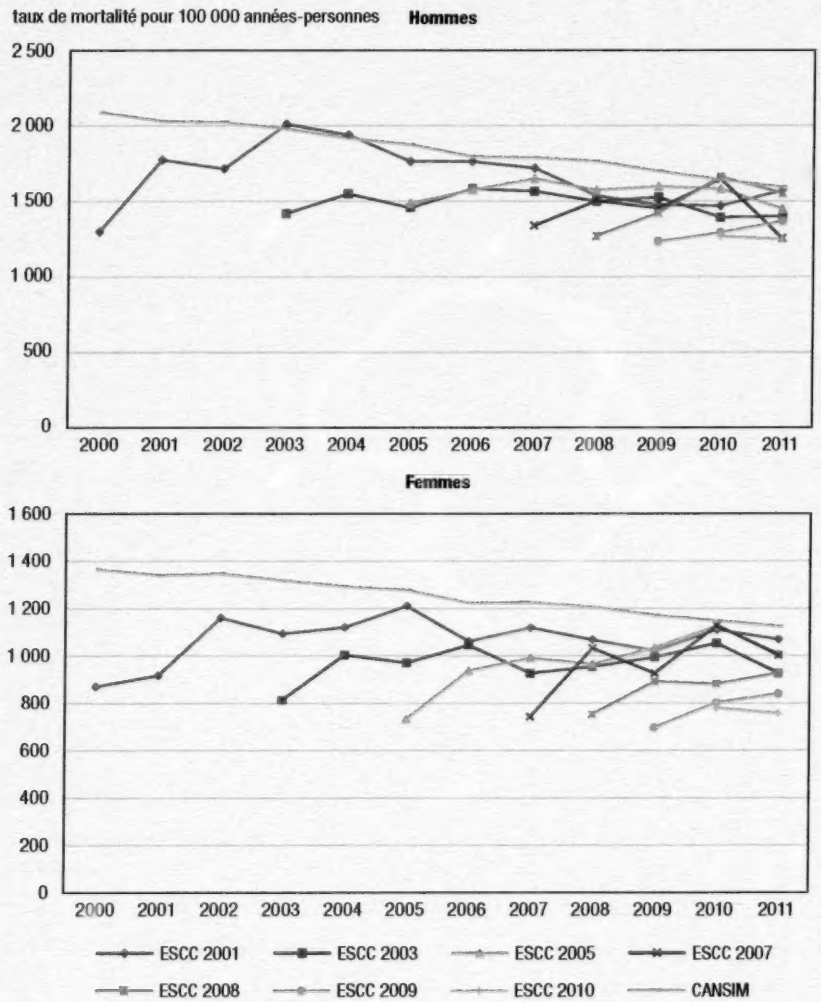
Les taux de faux positifs et de faux négatifs étaient de 0,04 % et 2,43 % respectivement (données non présentées dans les tableaux). Comme on pouvait s'y attendre, les taux de couplage (non pondérés) avec la BCDM étaient plus élevés chez les hommes (5,8 %) que chez les femmes (5,0 %) et augmentaient avec l'âge. Ils étaient les plus élevés chez les 75 ans et plus (27,0 %) (tableau 3). Ce profil de taux de couplage était similaire d'une année d'enquête à l'autre (données non présentées dans les tableaux).

Les répartitions géographiques des décès dans les données tirées du couplage entre l'ESCC et la BCDM étaient similaires à celles calculées à partir de la BCDM seulement. La majorité des

décès se sont produits dans les grandes provinces (Ontario, Québec et Colombie-Britannique) (tableau 4). Même si les tendances globales se sont maintenues, la répartition pondérée des décès d'une province et d'un territoire à l'autre se rapprochait davantage des résultats de la BCDM que les répartitions non pondérées. Les différences étaient évidentes particulièrement dans le cas des provinces de l'Atlantique.

La répartition des décès selon le groupe d'âge a révélé l'augmentation à laquelle on s'attendait aux âges avancés (tableau 4). Toutefois, en pourcentage de tous les décès, 49 % ont été enregistrés chez les personnes de 75 ans et plus dans les données couplées, comparativement à 61 % dans les données de la BCDM. La différence reflète en grande partie les décès non représentés dans les données de l'ESCC et de la BCDM couplées parmi les résidents d'établissements saisis dans la BCDM. Lorsque seuls étaient pris en compte les décès enregistrés dans la BCDM chez les personnes de 12 à 74 ans inclusivement, les répartitions étaient similaires (données non présentées). La répartition des décès selon la cause, établie à partir des données couplées, était similaire à la répartition fondée sur la BCDM (tableau 4).

Figure 1
Taux de mortalité normalisés selon l'âge pour 100 000 années-personnes, selon le sexe et l'année d'enquête, données couplées de l'Enquête sur la santé dans les collectivités canadiennes (ESCC) et de la Base canadienne de données sur la mortalité et de CANSIM, 2000 à 2011



Sources : Enquête sur la santé dans les collectivités canadiennes, 2000-2001 à 2011; tableau CANSIM 102-0504.

Les taux de mortalité annuels normalisés selon l'âge (TMNA) calculés à partir des données couplées de l'ESCC et de la BCDM étaient généralement moins élevés que ceux calculés à partir des taux de mortalité officiels (figure 1). Les TMNA étaient les plus faibles la première année de suivi pour chaque année de l'ESCC visée par l'étude, ce qui montre un biais possible dû à l'effet du « participant en santé », lequel constitue une raison à l'appui de la décision d'exclure la première année de suivi des modèles de survie.

Analyse de la survie

En général, l'analyse de la survie révèle des rapports de risques (RR) élevés pour la mortalité parmi les groupes les plus à risque (tableau 5). Les gros fumeurs (hommes = 2,36; femmes = 2,91) et les fumeurs légers (hommes = 1,92; femmes = 1,81) avaient des RR plus élevés comparativement aux non-fumeurs. Le risque de mortalité associé à l'IMC suivait une courbe en forme de U, les RR s'avérant élevés chez les personnes faisant partie de la catégorie d'IMC de l'insuffisance pondérale (hommes = 1,77; femmes = 1,50) et de celle de l'obésité de classe II et plus (hommes = 1,51; femmes = 1,20). Les RR pour les anciens buveurs (hommes = 1,65; femmes = 1,56) étaient élevés comparativement à ceux chez les personnes buvant modérément. Le fait de déclarer ne faire aucune activité physique ou en faire moins de 30 minutes par jour était associé à un risque de mortalité

plus élevé comparativement au fait d'en déclarer 60 minutes ou plus. Enfin, le fait de consommer moins de cinq portions de fruits et de légumes par jour était associé à un risque élevé de mortalité comparativement à plus de cinq portions. Les résultats associés à aucune portion de fruits et de légumes n'étaient pas significatifs, probablement en raison des petites tailles d'échantillon.

Discussion

Dans l'ensemble, 5,3 % des participants à l'Enquête sur la santé dans les collectivités canadiennes qui étaient admissibles au couplage ont fait l'objet d'un appariement à un enregistrement de décès dans la Base canadienne de données sur la mortalité. Les taux de faux positifs et de faux négatifs étaient de 0,04 % et 2,43 %, respectivement. L'utilisation du Fichier de données fiscales sommaires historiques a produit des taux de couplage avec la BCDM légèrement plus élevés. La validation externe a révélé dans les données couplées de l'ESCC et de la BCDM des profils de taux de mortalité qui étaient comparables aux estimations déclarées à l'échelle nationale. Ces profils attestent la qualité des données couplées et leur pertinence pour la recherche sur la mortalité au niveau de la population.

Les taux de mortalité normalisés selon l'âge calculés à partir des données couplées étaient uniformément plus faibles au cours des première et deuxième années de suivi, comparativement aux estimations nationales officielles de la mortalité. Cela était à prévoir, étant donné que les estimations officielles sont

Tableau 5
Rapports de risques corrigés selon l'âge pour la mortalité associée à certains comportements influant sur la santé, selon le sexe, participants de 20 ans et plus à l'Enquête sur la santé dans les collectivités canadiennes de 2003 et 2005 ayant fait l'objet d'un couplage à la Base canadienne de données sur la mortalité

	Hommes			Femmes		
		Intervalle de confiance de 95 %			Intervalle de confiance de 95 %	
Comportement influant sur la santé	Rapport de risques	de	à	Rapport de risques	de	à
Usage du tabac						
Non-fumeur[§]	1,00	...	...	1,00	...	...
Fumeur léger	1,92*	1,51	2,33	1,81*	1,52	2,11
Gros fumeur	2,36*	1,84	2,89	2,91*	1,92	3,91
Ancien fumeur	1,23*	1,05	1,40	1,31*	1,16	1,46
Indice de masse corporelle (IMC), avec correction						
Insuffisance pondérale (moins de 18,5)	1,77*	1,06	2,47	1,50*	1,16	1,85
Poids normal (18,5 à moins de 25,0)[§]	1,00	...	...	1,00	...	...
Embonpoint (25,0 à moins de 30,0)	0,87*	0,79	0,95	0,86*	0,78	0,94
Obésité de classe I (30,0 à moins de 35,0)	0,96	0,85	1,07	0,91	0,80	1,02
Obésité de classe II (35,0 ou plus)	1,51*	1,25	1,76	1,20*	1,00	1,40
Consommation d'alcool						
Buveur léger ou non-buveur	1,20*	1,09	1,31	1,15*	1,01	1,29
Personne buvant modérément[§]	1,00	...	...	1,00	...	...
Gros buveur	1,35*	1,06	1,63	2,29	0,41	4,16
Ancien buveur	1,65*	1,46	1,83	1,56*	1,36	1,75
Minutes d'activité physique par jour						
Aucune	1,89*	1,61	2,16	2,04*	1,65	2,43
Moins de 30	1,23*	1,08	1,38	1,22*	1,03	1,41
30 à 60	0,99	0,86	1,12	1,10	0,92	1,28
Plus de 60[§]	1,00	...	...	1,00	...	...
Portions de fruits et de légumes par jour						
Aucune	1,93	0,49	3,36	1,71	0,57	2,85
Moins de 2	1,52*	1,30	1,74	1,82*	1,51	2,13
2 à 5	1,18*	1,03	1,32	1,18*	1,05	1,32
Plus de 5[§]	1,00	...	...	1,00	...	...

... n'ayant pas lieu de figurer
* valeur significativement différente de celle pour la catégorie de référence ($p < 0,05$)
§ catégorie de référence
Sources : Enquête sur la santé dans les collectivités canadiennes, 2003 et 2005; Base canadienne de données sur la mortalité, 2004 à 2011.

Ce que l'on sait déjà sur le sujet

- Les registres de la statistique de l'état civil et les enquêtes nationales sur la santé sont des sources complémentaires de données.
- Lorsque les données administratives et les données d'enquête sur la santé sont combinées grâce au couplage d'enregistrements, les rapports entre les déterminants sociaux et les résultats pour la santé, comme le décès, peuvent être analysés de façon approfondie.
- Au Canada, les données sur la mortalité ont été couplées aux résultats du recensement pour examiner les différences de risque de décès entre groupes socioéconomiques, ainsi que pour calculer les taux de mortalité de sous-populations, comme les immigrants et les Autochtones.
- Une étude de l'Ontario a utilisé des données d'enquête sur la santé couplées à des données sur la mortalité pour explorer les associations avec l'usage du tabac, le régime alimentaire, l'activité physique et le stress.

Ce qu'apporte l'étude

- Les données de l'Enquête sur la santé dans les collectivités canadiennes (ESCC) de 2000-2001 à 2011 ont été couplées à celles de la Base canadienne de données sur la mortalité (BCDM).
- Dans l'ensemble, 5,3 % des participants à l'ESCC ont fait l'objet d'un couplage à un enregistrement de décès dans la BCDM.
- La similitude des profils de taux de mortalité fondés sur les données couplées et de ceux d'estimations déclarées à l'échelle nationale montre que les données couplées conviennent pour l'étude de la mortalité au niveau de la population au Canada.

fondées sur l'ensemble de la population, qui comprend les personnes vivant en établissement, tandis que les taux estimés à partir des données couplées de l'ESCC et de la BCDM représentent uniquement la population à domicile.

Des TMNA faibles au cours des deux premières années de suivi peuvent aussi être le résultat d'un biais associé à l'effet du « participant en santé ». Les participants à l'ESCC peuvent présenter un profil de morbidité et de mortalité plus favorable que les non-participants. Il se peut que les personnes malades ou en fin de vie soient moins susceptibles de répondre à une enquête. Les données d'une enquête écossaise sur la santé couplées à des données sur la mortalité ont révélé des taux de mortalité plus élevés parmi les non-participants (55 ans et plus) que les participants[34]. De même, en Finlande, des données d'enquête couplées avec des données sur la mortalité ont montré des taux de mortalité plus élevés chez les non-participants que chez les participants à l'enquête annuelle sur la santé[35].

Les différences dans les taux de mortalité pourraient tenir en partie à la situation démographique et socioéconomique des participants à l'enquête par rapport aux non-participants[35,36]. En outre, il se peut que les non-participants présentent un risque plus élevé de mortalité liée à l'alcool, à la drogue et à l'usage du tabac que les participants[37]. Cela porte à croire que la non-participation à l'enquête pourrait entraîner la sous-estimation de la mortalité globale et de l'incidence des comportements à risque.

Conformément aux résultats internationaux[18,38-41], l'analyse de la survie fondée sur les données couplées de l'ESCC et de la BCDM ont révélé des rapports de risques de mortalité plus élevés chez les fumeurs, les personnes ayant un IMC très faible ou très élevé, celles ayant déclaré moins de cinq portions de fruits et de légumes par jour, et celles ayant déclaré moins de 30 minutes d'activité physique par jour.

Pendant plus d'un demi-siècle, des études épidémiologiques importantes ont documenté les risques de mortalité prématurée liée à l'usage du tabac[42]. Une courbe en forme de U pour l'IMC et la mortalité figurait dans les recherches canadiennes et américaines[43]. Le profil de consommation d'alcool et de mortalité est conforme à une méta-analyse internationale[44], et le risque plus élevé parmi les anciens buveurs comparativement aux personnes buvant modérément appuie l'effet de l'« ex-buveur malade » (personnes qui arrêtent de boire en raison de l'apparition d'une maladie)[45].

Des recherches récentes[46,47] ont indiqué que, paradoxalement, certains facteurs de risque traditionnels peuvent s'inverser dans les faits aux âges avancés, hypothèse que l'on pourrait vérifier à partir des données couplées de l'ESCC et de la BCDM dans des analyses futures. D'autres déterminants sociaux de la santé, comme les interactions et le soutien social, qui peuvent contribuer au risque de mortalité, particulièrement chez les personnes âgées[48-51] pourraient être examinés à partir des données couplées.

Limites

Un niveau de prudence est justifié quant à l'utilisation des données couplées de l'ESCC et de la BCDM. Les faibles taux de mortalité la première année de suivi portent à croire à un biais attribuable à l'effet du « participant en santé ». Par conséquent, il est recommandé qu'au moins la première année de suivi soit exclue des estimations de la mortalité. De même, bien que l'ESCC soit une enquête transversale, il se peut que les personnes soient représentées plus d'une fois dans un cycle donné et d'un cycle à l'autre. Ces participants n'ont pas été supprimés, mais on les a plutôt identifiés et étiquetés, afin de les coupler au même enregistrement de décès. Toutefois, cela touche uniquement les analyses nécessitant plus d'une année d'enquête.

Mot de la fin

L'ESCC peut être couplée à la BCDM pour examiner les associations entre les facteurs de risque pour la santé et la mortalité. Le couplage représente une ressource pour la recherche sur les résultats au chapitre de la mortalité. Ces données peuvent servir à préciser les estimations des associations entre les comportements à risque et la mortalité dans des modèles prédictifs de la mortalité toutes causes confondues et de la mortalité par cause. ■

Remerciements

Les auteurs sont reconnaissants pour le financement reçu de Santé Canada, de l'Institut de recherche en services de santé/l'Institut de recherche de l'Hôpital d'Ottawa et de l'Université McGill pour appuyer le projet de couplage.

Références

1. C.D. Holman, A.J. Bass, D.L. Rosman *et al.*, « A decade of data linkage in Western Australia: Strategic design, applications and benefits of the WA data linkage system », *Australian Health Review*, 32(4), 2008, p. 766-777.
2. W.E. Winkler, « Record linkage », *Sample Surveys: Design, Methods and Applications*, 29A, 2009, p. 351-380.
3. M. Rotermann, G. Sanmartin, R. Trudeau *et al.*, « Couplage de données de recensement et de données d'hôpital », *Rapports sur la santé*, 25(10), 2014, p. 3-15.
4. K. Wilkins, M. Shields et M. Rotermann, « Utilisation des hôpitaux de soins de courte durée par les fumeurs – Une étude prospective », *Rapports sur la santé*, 20(4), 2009, p. 83-93.
5. R. Fransoo, P. Martens, H. Prior *et al.*, *Adult Obesity in Manitoba: Prevalence, Associations, and Outcomes*, Winnipeg, Manitoba, Manitoba Centre for Health Policy, 2011.
6. M. Davidsen, M. Kjoller et K. Helweg-Larsen, « The Danish National Cohort Study (DANCOS) », *Scandinavian Journal of Public Health*, 39(7 Suppl), 2011, p. 131-135.
7. P. Teckle, P. Hannaford et M. Sutton, « Is the health of people living in rural areas different from those in cities? Evidence from routine data linked with the Scottish Health Survey », *BMC Health Services Research*, 12, 2012, p. 43.
8. M. Hamer, G.D. Batty, E. Stamatakis et M. Kivimaki, « Comparison of risk factors for fatal stroke and ischemic heart disease: a prospective follow up of the health survey for England », *Atherosclerosis*, 219(2), 2011, p. 807-810.
9. R. Wilkins, M. Tjepkema, C. Mustard *et al.*, « Étude canadienne de suivi de la mortalité selon le recensement, 1991 à 2001 inclusivement », *Rapports sur la santé*, 19(3), 2008, p. 27-48.
10. S.W. Hwang, R. Wilkins, M. Tjepkema *et al.*, « Mortality among residents of shelters, rooming houses, and hotels in Canada: 11-year follow-up study », *British Medical Journal*, 339, 2009, b4036. doi:101136/bmj.b4036.
11. M. Tjepkema et R. Wilkins, « Espérance de vie restante à l'âge de 25 ans et la probabilité de survie jusqu'à l'âge de 75 ans, selon la situation socioéconomique et l'ascendance autochtone », *Rapports sur la santé*, 22(4), 2011, p. 31-36.
12. E. Ng, « Effet de l'immigrant en santé et taux de mortalité », *Rapports sur la santé*, 22(4), 2011, p. 25-30.
13. M. Tjepkema, R. Wilkins, S. Senécal *et al.*, « Années potentielles de vie perdus de 25 à 74 ans chez les Métis et les Indiens non inscrits, 1991 à 2001 », *Rapports sur la santé*, 22(1), 2011, p. 41-52.
14. J. Dankovchik, M.J. Hoopes, Warren-Mears et E. Knaster, « Disparities in life expectancy of Pacific Northwest American Indians and Alaska natives: analysis of linkage-corrected life tables », *Public Health Reports*, 130(1), 2015, p. 71-80.
15. A. Kariminia, T.G. Butler, S.P. Corben *et al.*, « Extreme cause-specific mortality in a cohort of adult prisoners–1988 to 2002: a data-linkage study », *International Journal of Epidemiology*, 36, 2007, p. 310-316.
16. S.A. Kinner, L. Degenhardt, C. Coffey *et al.*, « Substance use and risk of death in young offenders: A prospective data linkage study », *Drug and Alcohol Review*, 34, 2015, p. 46-50.
17. D. Nambiar, A. Weir, E.J. Aspinall *et al.*, « Mortality and cause of death in a cohort of people who had ever injected drugs in Glasgow: 1982-2012 », *Drug and Alcohol Dependence*, 1470, 2015, p. 215-221.
18. D.G. Manuel, R. Perez, C. Bennett *et al.*, *Sept années en plus : Incidence du tabagisme, de la consommation d'alcool, de l'alimentation, de l'activité physique et du stress sur la santé et l'espérance de vie en Ontario*, Un rapport de ICES/SPO, Toronto, Clinical Evaluative Sciences et Santé publique Ontario, 2012.
19. Statistique Canada, *Enquête sur la santé dans les collectivités canadiennes*, disponible à l'adresse http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3226&lang=en&db=imdb&adm=8&dis=2#b4
20. Statistique Canada, *Couplages d'enregistrements approuvés*, disponible à l'adresse http://www.statcan.gc.ca/fra/enregistrement/somm
21. Statistique Canada, *Directive sur le couplage d'enregistrements*, disponible à l'adresse http://www.statcan.gc.ca/fra/enregistrement/politique4-1
22. C.E. Shannon, « A mathematical theory of communication », *The Bell System Technical Journal*, 27, 1948, p. 379-423, 623-56.
23. H.B. Newcombe, « Record linking: The design of efficient systems for linking records into individual and family histories », *American Journal of Human Genetics*, 19, 1967, p. 335-359.
24. Statistique Canada, *Summary Report of the Canadian Community Health Survey (CCHS) – Canadian Mortality Database (CMDB) Record Linkage Study* (document interne), Ottawa, Statistique Canada, Division de la statistique de la santé, 2015.
25. I.P. Fellegi et A.B. Sunter, « A theory for record linkage », *Journal of the American Statistical Association*, 64(328), 1969, p. 1183-1210.
26. Statistique Canada, *G-Coup Version 3.0 guide de l'utilisateur* (document interne).
27. M. Rotermann, C. Sanmartin, R. Trudeau et H. St-Jean, « Couplage des données du Recensement de 2006 et des données hospitalières au Canada », *Rapports sur la santé*, 26(10), 2015, p. 11-23.
28. Agency for Healthcare Research and Quality, *Linking Data for Health Services Research: A Framework and Instructional Guide*. AHRQ Publication No. 14-EHC033 EF. Rockville, Maryland, Department of Health and Human Services, 2014.
29. I. Dohoo, W. Martin et H. Stryhn, *Methods in Epidemiologic Research*, Charlottetown, Prince Edward Island: VER Inc., 2012.
30. Statistique Canada, *Tendances dans l'évolution des taux de mortalité, 2000 à 2011*, disponible à l'adresse http://www.statcan.gc.ca/pub/82-625-x/2014001/article/11897-fra.htm
31. Statistique Canada, Décès, selon l'âge et le sexe, Canada, Tableau 102-0503 CANSIM, disponible à l'adresse http://www5.statcan.gc.ca/cansim/a26
32. Statistique Canada, Décès et taux de mortalité, selon le groupe d'âge et le sexe, Canada, provinces et territoires, Tableau 102-0504 CANSIM, disponible à l'adresse http://www5.statcan.gc.ca/cansim/pick-choisir?lang=fra&p2=33&id=1020504
33. M. Shields, S. Connor Gorber, I. Janssen et M.S. Tremblay, « Biais dans les estimations autodéclarées de l'obésité dans les enquêtes canadiennes sur la santé : le point sur les équations de correction applicables aux adultes », *Rapports sur la santé*, 22 (3), 2011, p. 39-50.
34. G. Cohen et J.C. Dugg, « Are nonrespondents to health surveys less healthy than respondents? », *Journal of Official Statistics*, 18(1), 2002, p. 13-23.
35. H. Tolonen , T. Laatikainen, S. Helakorpi *et al.*, « Marital status, educational level and household income explain part of the excess mortality of survey non-respondents », *European Journal of Epidemiology*, 25(2), 2010, p. 69-76, doi: 10.1007/s10654-009-9389-9. Epub 2009 Sep 25.
36. M. Laaksonen, K. Talala, T. Martelin *et al.*, « Health behaviours as explanations for educational level differences in cardiovascular and all-cause mortality: a follow-up of 60,000 men and women over 23 years », *European Journal of Public Health*, 18(1), 2007, p. 38-43.

37. A.L. Christensen, O. Ekholm, L. Gray *et al.*, « What is wrong with non-respondents? Alcohol, drug- and smoking related mortality and morbidity in a 12-year follow up study of respondents and non-respondents in the Danish Health and Morbidity Survey », *Addiction*, 110(9), 2015, p. 1505-1512.

38. E.S. Ford, G. Zhao, J. Tsai et C. Li, « Low-risk lifestyle behaviors and all-cause mortality: Findings from the National Health and Nutrition Examination Survey III Mortality Study », *American Journal of Public Health*, 101(10), 2011, p. 1922.

39. E.S. Ford, M.M. Bergmann, H. Boeing *et al.*, « Healthy lifestyle behaviors and all-cause mortality among adults in the United States », *Preventive Medicine*, 55(1), 2012, p. 23-27.

40. G. Behrens, B. Fischer, S. Kohler *et al.*, « Healthy lifestyle behaviors and decreased risk of mortality in a large prospective study of U.S. women and men », *European Journal of Epidemiology*, 28(5), 2013, p. 361-372.

41. O. Oyebode, V. Gordon-Dseagu, A. Walker et J.S. Mindell, « Fruit and vegetable consumption and all-cause, cancer and CVD mortality: Analysis of Health Survey for England data », *Journal of Epidemiology and Community Health*, 68(9), 2014, p. 856-862.

42. M.J. Thun, « Public Health Classics: When truth is unwelcome: The first reports on smoking and lung cancer », *Bulletin of the World Health Organization*, 83(2), 2005, p. 144-145.

43. K.M. Flegal, B.K. Kit, H. Orpana et B.I. Graubard, « Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis », *Journal of the American Medical Association*, 309(1), 2013, p. 71-82.

44. P.E. Ronksley, S.E. Brien, B.J. Turner *et al.*, « Association of alcohol consumption with selected cardiovascular disease outcomes: A systematic review and meta-analysis », *British Medical Journal*, 342 (7795), 2011.

45. A.G. Shaper, « Alcohol consumption decreases with the development of disease », *Addiction*, 106(5), 2011, p. 1023-1025.

46. R.S. Peter, B. Mayer et H. Concin, « The effect of age on the shape of the BMI-mortality relation and BMI associated with minimum all-cause mortality in a large Austrian cohort », *International Journal of Obesity*, 39(3), 2015, p. 530-534.

47. S.F. Ahmadi, E. Streja et G. Zahmatkesh, « Reverse epidemiology of traditional cardiovascular risk factors in the geriatric population », *Journal of the American Medical Directors Association*, 16(11),2015, p. 933-939.

48. Centers for Disease Control, *Heat Waves 2009*, [cited August 1 2014], disponible à l'adresse http://www.cdc.gov/climateandhealth/effects/heat.htm. document consulté le 1er août 2014.

49. E. Kleninberg, « Dying alone: The social production of urban isolation », *Ethnography*, 2(4), 2001, p. 501-531.

50. J. Holt-Lunstad, T.B. Smith et J.B. Layton, « Social relationships and mortality risk: a meta-analytic review », *PLoS Medicine*, 7(7), 2010, p. 1000316.

51. S. Stringhini, L. Berkman, A. Dugravot *et al.*, « Socioeconomic status, structural and functional measures of social support, and mortality: The British Whitehall II Cohort Study, 1985-2009 », *American Journal of Epidemiology*, 17 (12), 2012, p. 1275-1283.